

# Lessons Learned from Usability Tests with a Collaborative Cognitive Workspace for Human-Robot Teams \*

Julie L. Marble, David J. Bruemmer, Douglas A. Few

Human, Robot and Remote Systems  
Idaho National Engineering  
and Environmental Laboratory  
Idaho Fall, ID  
{marbjl, bruedj, fewda}@inel.gov

**Abstract** - *Use of new robotics technologies is challenged by issues of system trust, unknowns regarding how the system will, can, and should be used, and possibilities for human error that may cause harm to the human operator, system, or environment. This paper discusses initial usability tests of a mixed-initiative robotic system. Participants were asked to search a building using a robot equipped with multiple levels of autonomy to identify 3 targets in pre-specified locations. The experiment showed a significant difference between novice and experienced robotic operators especially regarding willingness to use the autonomous capabilities of the robot. Users unfamiliar with teleoperation were more willing to utilize the autonomous capabilities of the robot, while skilled teleoperators preferred and were more efficient when in direct control. Users were almost always able to successfully complete the search task. However, feedback indicates that users, having been given only a cursory explanation of the system, were sometimes confused by robot initiative even though the interface supplied textual explanations. The experiment shows that mixed-initiative interaction may exceed the limitations of either fully autonomous or teleoperated control; however, potential benefits can easily be overshadowed by control challenges inherent to deploying robot – human teams.*

**Keywords:** Mixed-Initiative Control, Dynamic Autonomy, Human-Robot Interaction, Robots

## 1 Introduction

Remote sensor deployment is an application area where robots can provide benefits in terms of time, cost, safety, and quality of data [2]. However, these benefits can be realized only if this technology is willingly accepted by the users, and designed to meet their needs [11, 12]. Previously, the INEEL successfully deployed a teleoperated robotic system

to characterize an area that had been closed to human entry for many years. Although the 2001 deployment was a success in terms of human exposure, time, and money, evaluation of the operation exposed severe limitations to the master-slave strategy employed, including lapses in communication and situation awareness that forced personnel to enter the environment [3]. Consequently, the INEEL has since developed a mixed-initiative command and control architecture, which has been implemented on a variety of ground vehicles ranging from 25 to 300lbs. This paper discusses the findings of recent human participant usability testing intended to evaluate the efficacy of the robot's autonomous behaviors, the utility of various interface components and also the operators' willingness to permit and exploit robot initiative.

Human Factors studies in the area of human-computer interaction (HCI) and human-machine interaction (HMI) have revealed that many complex tasks are more successfully performed when the system is designed to support the needs of the human rather than eliminating the human from the system (e.g., [1, 5]). In fact, in many cases, the goal to eliminate the human from the system has resulted in significant system failures and even human deaths specifically because the system was not designed to support interaction with the human (see [4, 5] for multiple examples). Rather than eliminate the human from the loop the control architecture used in this study supports different levels of human input.

Within the field of robotics, both fully autonomous approaches and teleoperated approaches have failed to realize the immense potential for humans and robots to work together as a true team. The fundamental aspect of a team that distinguishes it from a simple group is the presence of a shared

---

\* 0-7803-7952-7/03/\$17.00 © 2003 IEEE.

goal. Effective teams typically cooperate and anticipate the needs of teammates via a shared mental model of the task and current situation [6, 7, 13]. The cognitive workspace used in this usability study provides a means to represent this common goal with a common form of representation that is meaningful to both robot and human. Using this workspace, this paper considers the potential for human-robot teams to move beyond the supervisory paradigm that has dominated thinking about autonomous control (e.g., [8, 9, 10]), towards peer-peer interaction where each member can actively and authoritatively take initiative to accomplish task objectives.

## 2 Experimental Design

### 2.1 Participants.

Eleven INEEL employees from the pool of INEEL employees participated. Seven participants had little or no prior experience with remote or robotic systems or some experience. Four expert users also participated, defined by extensive (average 7 years) primarily job related experience using teleoperated vehicles, master-slave systems, or similar systems.

### 2.2 Test Location

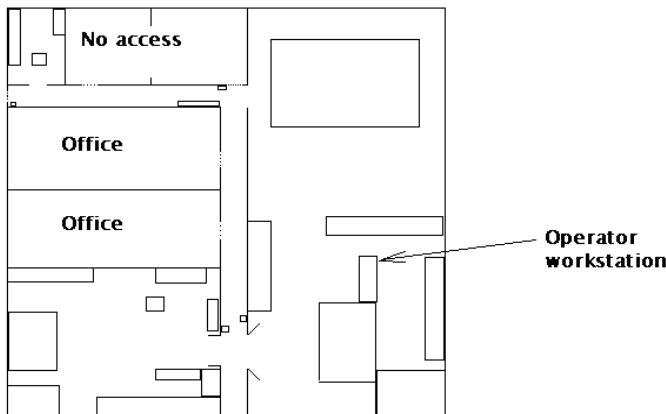


Figure 1. Schematic representation of location of testing.

The layout of the building used for testing is shown schematically in Figure 1. The building is made of steel and concrete, which disrupts video feed signals, as could often be encountered in a “real world” setting. The building is approximately 3600 square feet, 60 feet long by 60 feet wide. The central hallway is approximately 5 feet wide and 40 feet long; the distance from the left turn to the building

exit is approximately 30 feet. This turn was narrowed by coveralls and boots placed in a hanging area. Opposite these, a box (18” x 12”) was placed next to the doorway, depicted in the schematic, to further narrow the hallway. Two targets were located in opposing corners of the storage room, in areas that required the robot to be maneuvered past obstacles in the room in order to identify the targets. Just outside the storage room doors, the hallway was narrowed by two trash cans placed 4.25 feet diagonally from one another, as shown in the schematic. The third primary target was located in the janitorial room at the end of the central hallway. Operators were asked to not allow the robot to enter the offices.

### 2.3 System Design



Figure 2. The ATRV Jr. and component sensors

The robot used in these tests was an ATRVJr (see Figure 2), with a control architecture developed at the INEEL. The robot is equipped with a variety of range sensor information including inertial sensors, compass, wheel encoders, laser range finders, computer vision, infrared break beams, tilt sensors, bump sensors, and sonar. Streaming video from the robot is provided to the operator. The robot abstracts information about the environment at many levels including terse textual descriptions of the robot’s local surroundings and the choices (depending on the level of autonomy) that face the human user. In addition, the robot uses simultaneous mapping and localization to build a representation of the environment on the fly (shown in the upper right quadrant of Figure 3).

The architecture controlling the robot supports five levels of human-robot interaction:

**Teleoperation:** The user has full, continuous control of the robot at a low level. The robot takes no

initiative except to stop once it recognizes that communications have failed. It does indicate the detection of obstacles in its path to the user, but will not prevent collision.

**Safe Mode:** User directs movements of robot, but the robot takes initiative and has the authority to protect itself based on its proprioception and self-status evaluation; for example, it will stop or refuse to move before it collides with an obstacle that it detects via multiple sensors.

**Shared Control:** The robot takes the initiative to choose its own path in response to general direction input from the operator. Although the robot handles the low level navigation and obstacle avoidance, the user supplies intermittent input to guide the robot in general directions.

**Full Autonomy:** (Although this mode was available, it was not used during the test.) The robot navigates, requiring no operator input except high-level tasking such as "follow that target" or "search this area."

**Dynamic Control.** The user is able to switch between any of the above modes of autonomy to best accomplish the task by selecting the mode from the touch screen.



Figure 3. Current user interface.

For each level of autonomy, data are fused into a specialized interface presented on a touch screen (shown in Figure 3) that provides the user with abstracted graphical and textual representations of the environment and task appropriate for the current mode. Immediate obstacles that inhibit motion are shown as red ovals to the side or to the front or back and resistance to motion is shown with arcs pulsing from the wheels in the iconographic representation at

the bottom right of Figure 3. Video information is provided to the operator continuously. As the operator touches the video display, the robot's camera aligns to center that part of the image. The operator can also manipulate the camera by selecting the tilt and pan buttons located around the video display. The robot relays high-level information to the user in text form using the feedback textbox below the image window. The robot status window (lower left of Figure 3) provides a variety of information including pitch and roll, power, heading, speed and a fusion of this information into a single measurement of "health." The user may direct the robot by touching the arrows or using a joystick, both of which were available during the test.

## 2.4 Method

Participants were asked to search a building to locate and identify 3 targets in pre-specified locations as quickly and safely as possible. Targets were selected at random from a set of 20 different stuffed animals varying in size from 6 inches tall to approximately 1 foot tall. Participants made four searches of the building; in three searches, participants were limited to only one level of autonomy (teleoperation, safe teleoperation, or shared autonomy) and the order of use was randomized, while in the fourth search, participants were allowed to shift the level of autonomy as much as they desired to accomplish the task (dynamic autonomy).

During each search, two alternate targets were placed along the route the robot would need to navigate, although the locations of these secondary targets were randomized on each trial. Participants were told to be aware that up to two of these secondary targets might be present, but that they were not required to find them. These targets were used to assess the degree of situation awareness participants had about the environment of the robot while engaged in navigation.

During performance of the search task, the joystick, touch screen, and laptop were videotaped to record the actions of the operator. After each search, the participant was asked to fill out a 19 item subjective assessment regarding the task just completed, rating items from 1 ("very true") to 5 ("very false"). This form is available upon request. Finally, a camera person followed the robot throughout the building to record the actions of the robot and to prevent the robot from damaging walls – a serious concern especially in teleoperation mode.

Before beginning the searches, participants were given 20 minutes to familiarize themselves with the interface and behavior of the robot. During this time, participants were asked to perform simple tasks using each level of autonomy.

### 3 Results and Discussion

At the start of the experiment, participants experienced with tethered robotic systems often spent the first few minutes interacting with the robot adjusting the camera down and panning around the floor. Interviews with these participants revealed that they were attempting to learn how much of the robot could be seen in the camera, in order to determine how far forward they could see; that is to get an indication of depth from the video feed. Relative to this, most participants indicated a desire for the interface to overlay the video with a depth indicator, especially in teleoperated mode. Analysis of the videotapes supports this conclusion as well. In addition, several participants had difficulty moving through doorways in all modes because they were attempting to start turns up to three feet too early, and therefore, encountering the obstacles. In addition, several participants were very jerky in their control of the robot, inching the robot forward towards and around obstacles. This tedious, inefficient control was usually due to operators that focused on the video display. Using the video alone makes it very difficult to discern where the robot is relative to obstacles in the environment.

In contrast, some participants quickly learned to navigate around obstacles, not based on the video information, but rather on the obstacle indicators on the interface. Surprisingly, these operators used the abstracted representations from the robot to the almost complete neglect of the video information, until they had reached space in which none of the obstacle indicators were lit. Notably, the participant who was most efficient with this technique had no previous robotic experience, and discovered this technique on her own. Several participants were highly successful and very quick in navigating corners and the slalom based entirely on this information, while others did not appear to learn its utility or importance. It is not clear at this time why not all participants were able to maneuver in this manner.

Adaptation to greater levels of autonomy varied greatly across the participants, but was statistically independent of previous experience or skill. In the

shared mode, participants who had reported no previous experience using remote systems rated as more false that « predicting the outcome of control use was difficult » than participants who had previous experience using remote systems,  $F(1, 11) = 6.364$ ,  $MSE = 3.0$ ,  $p < .04$ . Participants with no remote system experience rated this question as false on average (mean rating = 4), while those with experience rated it as true on average (mean rating = 2). Participants with the most previous experience using remote systems also tended to rate the controls as more difficult to predict in the teleoperated sessions than did the participants who had no previous experience with remote systems,  $F(1, 11) = 6.000$ ,  $p = .07$ . This was modified by a significant interaction with session,  $F(1, 11) = 16.667$ ,  $MSE = 2.083$ ,  $p < .015$ , which indicated that experienced participants rated as less true that « prediction was much more easy in session 1 than in session 2 » (4.0 [false] versus 1.0 [true]), while participants with no experience rated as more true that prediction of the controls was easy in sessions 1 and 2 (1.5 versus 2.0 [between very true and true], respectively). This may indicate that because inexperienced participants had no previous expectations regarding how the system would work, they were better able to learn how to interact with the robot without training.

Two participants were most notably able to adapt to the use of shared control. Interestingly, both of these participants had no previous experience with remote systems. In addition, one of these two participants had the most difficulty with the teleoperated mode, as indicated by the number of physical contacts and intercessions needed during this session (4 contacts where intercession was required compared to a mean of 0.4 for all other participants). This participant was also the oldest participant whose previous work experience included 17 years as a hydraulic backhoe operator. Interviews with this participant revealed that his difficulty in teleoperation mode arose from the fact that he was used to experiencing force feedback from the controls. The second of these two participants tended to be very hesitant issuing commands to the robot, especially in teleoperated mode, often issuing a command then waiting to see how far the robot would get before making another command with the joystick. Of all participants, these two were most likely to physically release the controls in shared mode after making a command.

Participants with the most experience with teleoperated systems reported the most frustration in

shared mode. Analysis of the videos indicated that these participants attempted to give constant commands to the robot in shared mode, just as in teleoperated mode. The difficulty these participants experienced arose from the robot attempting to take initiative to correct its course away from an obstacle in a direction contrary to that being commanded. Because these participants were also less likely to navigate based on the obstacle display than less experienced participants, they were also less likely to realize that the robot was protecting itself from collisions by correcting its course. In addition, few if any users noted the textual information indicating obstacles located below the video feed. This finding indicates that interfaces to support human-robot interaction must use highly salient cues to indicate why a robot is refusing a command, and that these cues must be emphasized more when users have previous remote experience.

Participants tended to be more certain of having identified the secondary targets in safe mode than in other modes, although this may be due to the fact that shared control allowed operators to devote less time to the video display. Most operators were able to identify at least one of the secondary targets, which implies that the information presented in the interface was sufficient to allow the participants to maintain situation awareness beyond the specific task structure (i.e. speed of performance for finding the primary targets) which was emphasized in the experiment.

Mean certainty ratings of spotting the secondary targets were analyzed with respect to the actual number of secondary targets identified. Certainty of having identified all the secondary targets was a significant, positive indicator of actual performance finding the secondary targets,  $F(4, 31) = 3.469$ ,  $MSE = 1.640$ ,  $p < 0.022$ . In other words, the more certain participants were that they had identified all secondary targets, the more likely they were to have actually identified all secondary targets.

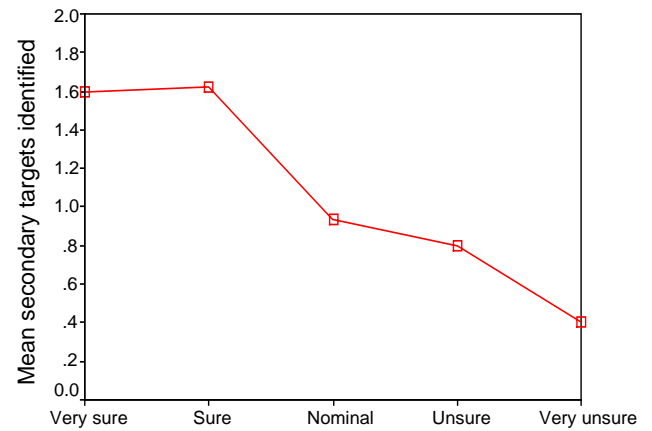


Figure 4. Number of actual secondary targets identified as a function of certainty.

## 4 Conclusions

The usability tests revealed surprising results regarding the manner in which human and robot team members can and should interact. Perhaps most surprising was the disparity between users. Some users preferred utilizing the autonomous capabilities of the robot and wanted to learn more about them, while others preferred the teleoperated approach precisely because it required less understanding. Some operators utilized multiple modalities presented within the interface such as mapping, vision and abstracted range sensing while others seemed unable to process this much information and often focused entirely on the video display. Critical to the design of autonomous systems, this usability test revealed that the participants who performed best in teleoperated and/or safe modes were not the participants who were best able to adapt to the shared control structure. There were several indicators that previous experience with remote systems made learning interaction with an autonomous system more difficult.

Future experiments will provide at least some level of training for the participants which may greatly increase the participants' understanding of how and why the robot initiative kicks in. Although reduction in training time has often been given as a reason for developing robot intelligence, it may well be that the higher the level of robot autonomy, the more training is needed by the human operator to effectively understand how and why the robot will behave the way it does. While the usability study has illuminated many opportunities for improvement, we

believe that already the collaborative, cognitive workspace offers a means to mediate between human and robot team members, providing a means to fuse sensing from differing modalities and communicate knowledge from disparate perspectives.

## 5 References

- [1] K. A. Abbott, S. M. Slotte, & D. K. Stimson, Federal Aviation Administration Human Factors Team Report on: The Interfaces Between Flightcrews and Modern Flight Deck Systems, Federal Aviation Administration, Washington, DC, Tech. Rep. 1996. [Online]. Available: <http://www.faa.gov/avr/afs/interfac.pdf>.
- [2] D. J. Bruemmer, J. L. Marble, M. O. Anderson, M. D. McKay, D. D. Dudenhoeffer. Dynamic-Autonomy for Remote Robotic Sensor Deployment, Proc. Spectrum 2002, Reno, NV, August 2002.
- [3] D. J. Bruemmer, J. L. Marble, D. D. Dudenhoeffer, M. O. Anderson, M. D. McKay. "Mixed-Initiative Control for Remote Characterization of Hazardous Environments." Proc. HICSS 2003, Waikoloa Village, Hawaii, January 2003.
- [4] S. Casey, Set Phasers on Stun: And Other True Tales of Design, Technology, and Human Error (2nd edition). Santa Barbara, CA: Aegean Publishing Co. 1998.
- [5] J. A. Espinosa, J. Cadiz, L. Rico-Gutierrez, R. E. Kraut, W. Scherlis, & G. Lautenbacher, "Coming to the wrong decision quickly: Why awareness tools must be matched with appropriate tasks." Proc. of the Human Factors in Computing Systems, Computer-Human Interaction (CHI) conference, ACM Press, 2000.
- [6] J. A. Espinosa, R. Kraut, J. Lerch, S. Slaughter, H. Herbsleb, & A. Mockus, "Shared mental models and coordination in large-scale, distributed software development." Proc. of the Twenty-Second International Conference on Information Systems, (ICIS) 2001, New Orleans, LA.
- [7] J. L. Marble, D. Bruemmer, D. Few, & D. Dudenhoeffer, "Evaluation of supervisory vs. peer-peer interaction for human-robot teams." To appear in Proc. of the 37th Annual Hawaii International Conference on System Sciences, Waikoloa Village, Hawaii, submitted.
- [8] T. B. Sheridan (1984). Supervisory control of remote manipulators, vehicles and dynamic processes: Experiments in command and display aiding. In W. Rouse, (Ed.) *Advances in Man-Machine Systems Research* (Vol. 1), Greenwich, Connecticut: JAI Press Inc.
- [9] T. B. Sheridan (1997). Supervisory control. In G. Salvendy (Ed.), *Handbook of Human Factors* (2nd ed.) NY, NY: Wiley, pp. 1295 - 1327.
- [10] T. B. Sheridan & W. L. Verplank (1978). Human and computer control of undersea teleoperators. Man-Machine Systems Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass., Tech. Report.
- [11] C. D. Wickens (1992). *Engineering Psychology and Human Performance* (2nd ed.), NY, NY: Harpers Collins Publishers, Inc.
- [12] C. D. Wickens, S. Gordon, & Y. Liu (1998). *An Introduction to Human Factors Engineering*. New York: Addison-Wesley-Longman.
- [13] J. Yen, J. Yin, T. R. Ioerger, M. Miller, D. Xu, & R. Volz, "CAST: Collaborative Agents for Simulating Teamwork." In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI '2001)*, Seattle, WA.